The effect of vitamin C on tooth growth in guinea pigs

José Rodrigo Cervantes Polanco

Overview

In this project we are going to analize the length of odontoblasts (teeth) in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods (orange juice or ascorbic acid). We are going to make a summary, some basic exploratory data analysis and a basic inferential data analysys.

Basic settings

We fixed a few basic settings to be used during this analysis.

```
setwd("~/Documents/ProgramsGitHub/StatInference")  # working directory
library(ggplot2)      # to be able to graph
data(ToothGrowth)     # loading the data
```

Summary

Here we are going to do a summary of our data by checking the information given in the overiew. We have a data frame with 60 observations on 3 variables: len:lenght of teeth, supp:delivery method and dose:dose level of vitamin C.

```
str(ToothGrowth)
```

```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
summary(ToothGrowth$len)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     4.2    13.1    19.2    18.8    25.3    33.9
```

```
levels(ToothGrowth$supp)
```

```
## [1] "OJ" "VC"
```

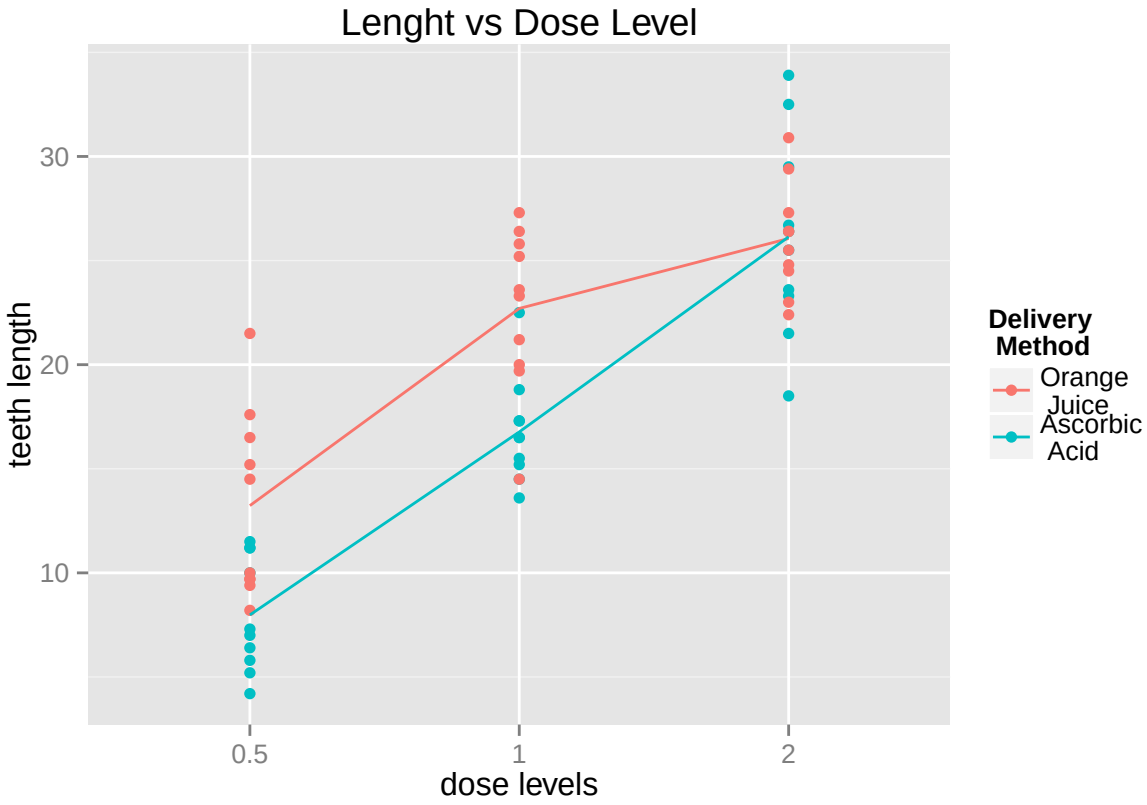
```
levels(as.factor(ToothGrowth$dose))
```

```
## [1] "0.5" "1"   "2"
```

Basic exploratory data analysis

We graph the data by teeth lenght versus dose levels for each kind of delivery method

```
g <- ggplot(data=ToothGrowth,aes(x=factor(dose),y=len, group=supp, colour=supp))
g <- g + geom_point() + stat_smooth(se=FALSE)
g <- g + scale_color_discrete("Delivery \n Method",breaks=c("OJ","VC"),
                              labels=c("Orange \n Juice","Ascorbic \n Acid"))
g <- g + labs(title="Lenght vs Dose Level",x="dose levels",y="teeth length")
g
```



It seems that the orange juice delivery method is more effective that ascorbic acid delivery method for doses of 0.5 and 1 mg, but for doses of 2 mg both delivery methods have similar effect.

Basic inferential data analysys

We will examine the predictions from previous section by using hypothesis tests analysis.

For this project we are going to assume that the data is I.I.D. normal, there is no pairing and there is difference variance per group (because of lack of information from the dataset).

We run a t-test in the data restricted to 0.5 mg dose with different delivery method, our null-hypothesis is: H_0 : the difference of the means of the two data sets with different delivery method is zero.

```
t.test(ToothGrowth[ToothGrowth$supp=="OJ" & ToothGrowth$dose ==0.5,]$len,
       ToothGrowth[ToothGrowth$supp=="VC" & ToothGrowth$dose ==0.5,]$len,
       paired = F, var.equal = F)
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 0.5,  and ToothGrowth[ToothGrowth$s
## t = 3.17, df = 14.97, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719 8.781
## sample estimates:
## mean of x mean of y
##     13.23      7.98
```

We have enough evidence to reject the null hypothesis and conclude that the mean of the 0.5 dose delivered by orange juice were significantly different from the mean of the 0.5 dose delivered by ascorbic acid. When substracting (orange juice - ascobic acid) the interval is entirely above zero; the orange juice appears to be a more effective delivery method.

We run a t-test in the data restricted to 1 mg dose with different delivery method, our null-hypothesis is: H_0 : the difference of the means of the two data sets with different delivery method is zero.

```
t.test(ToothGrowth[ToothGrowth$supp=="OJ" & ToothGrowth$dose ==1,]$len,
       ToothGrowth[ToothGrowth$supp=="VC" & ToothGrowth$dose ==1,]$len,
       paired = F, var.equal = F)
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 1,  and ToothGrowth[ToothGrowth$su
## t = 4.033, df = 15.36, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802 9.058
## sample estimates:
## mean of x mean of y
##     22.70     16.77
```

We have enough evidence to reject the null hypothesis and conclude that the mean of the 1 dose delivered by orange juice were significantly different from the mean of the 1 dose delivered by ascorbic acid. When substracting (orange juice - ascobic acid) the confidence interval is entirely above zero; the orange juice appears to be a more effective delivery method.

We run a t-test in the data restricted to 2 mg dose with different delivery method, our null-hypothesis is: H_0 : the difference of the means of the two data sets with different delivery method is zero.

```
t.test(ToothGrowth[ToothGrowth$supp=="OJ" & ToothGrowth$dose ==2,]$len,
       ToothGrowth[ToothGrowth$supp=="VC" & ToothGrowth$dose ==2,]$len,
       paired = F, var.equal = F)
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 2,  and ToothGrowth[ToothGrowth$su
## t = -0.0461, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.798  3.638
## sample estimates:
## mean of x mean of y
##     26.06     26.14
```

We don't have enough evidence to reject the null hypothesis and we can not conclude that there is a significant difference between the mean of the 2 dose delivered by orange juice and the mean of the 2 dose delivered by ascorbic acid. When substracting (orange juice - ascobic acid) the confidence interval contains zero; there is not evidence that the orange juice is more effective that ascobic acid.